The Impact of High-Speed Railway on Firm Entry in China

Jidong Ma

A Dissertation
Presented to the Faculty
of Cornell University
in Candidacy for the Degree
of Master of Science

August 2018

Abstract

Firm entry has been an important engine for economic growth in China. The expansion of high-speed railway (HSR) system holds the promise of helping firm entry by reducing the transportation cost and facilitating the flows of human capital and knowledge. In this thesis, I conduct an empirical analysis to quantify the impact of HSR expansion on firm entry in China. The data consists of detailed firm registration data at the county level from 2011 to 2015 and HSR coverage information during the same period. I use the difference-in-differences approach to estimate the impact of HSR expansion on firm entry. I find that the opening of an HSR station in a county increases firm entry in that county by 5.72% on average and that the impact is larger for high-tech industries. The analysis suggest that the improvement in transportation infrastructure such as the HSR expansion could promote entrepreneurship and further economic growth in China.

Biosketch

Jidong Ma was born in Dongying, China in 1992. He received his Bachelor's degree from the Department of Sociology at Tsinghua University in 2015. He joined Cornell University in 2016 to pursue his Master's degree. During his graduate study in Ithaca, he focused his research on transportation infrastructure investment, firm activities and economic development in China.

Acknowledgements

I would like to offer my deep gratitude to my advisor, Professor Shanjun Li, for his patience and support throughout my graduate life at Cornell University. His guidance and attitude toward research have been constant aspirations, for which I will always be grateful.

I would also thank Professor Kieran Donaghy for his advice on this research and his insights on regional economic development in general. His genuine care for student learning is truly inspirational. I will try my best to not let him down in my future career.

Thanks also go to Ms. Ruyu Chen for her guidance and to all my friends in Ithaca for sharing blizzards and sunset glow together.

To my beloved family.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Large-scale infrastructure projects such as roads and waterways have played important roles in the social and economic development throughout human history. These projects often involves a large amount of funding and are public in nature. At the same time, entrepreneurship is another critical element for economic development. Whether and how these types of public investment on infrastructure have shaped entrepreneurial activities are interesting and important questions that speak directly to the benefit of these investment. In this thesis, I examine these questions by leveraging the rapid expansion of the High-Speed Railway (HSR) network in China.

China's recent efforts to expand railway construction are largely driven by the increasing travel demand from continued economic growth. The economy of China has been growing rapidly over past nearly 40 years since the economic reforms in 1978. After the turn of the 21^{st} century, China's economy surpassed France, UK, Germany and Japan and became the worlds second largest. Large infrastructure investment by Chinas government played a key role in this process. As urbanization intensifies, China continues to invest in High Speed Railway (HSR) to fulfill the emerging new transport needs and rebalance the growing regional imparity to reduce poverty and achieve an inclusive and sustainable economic growth.

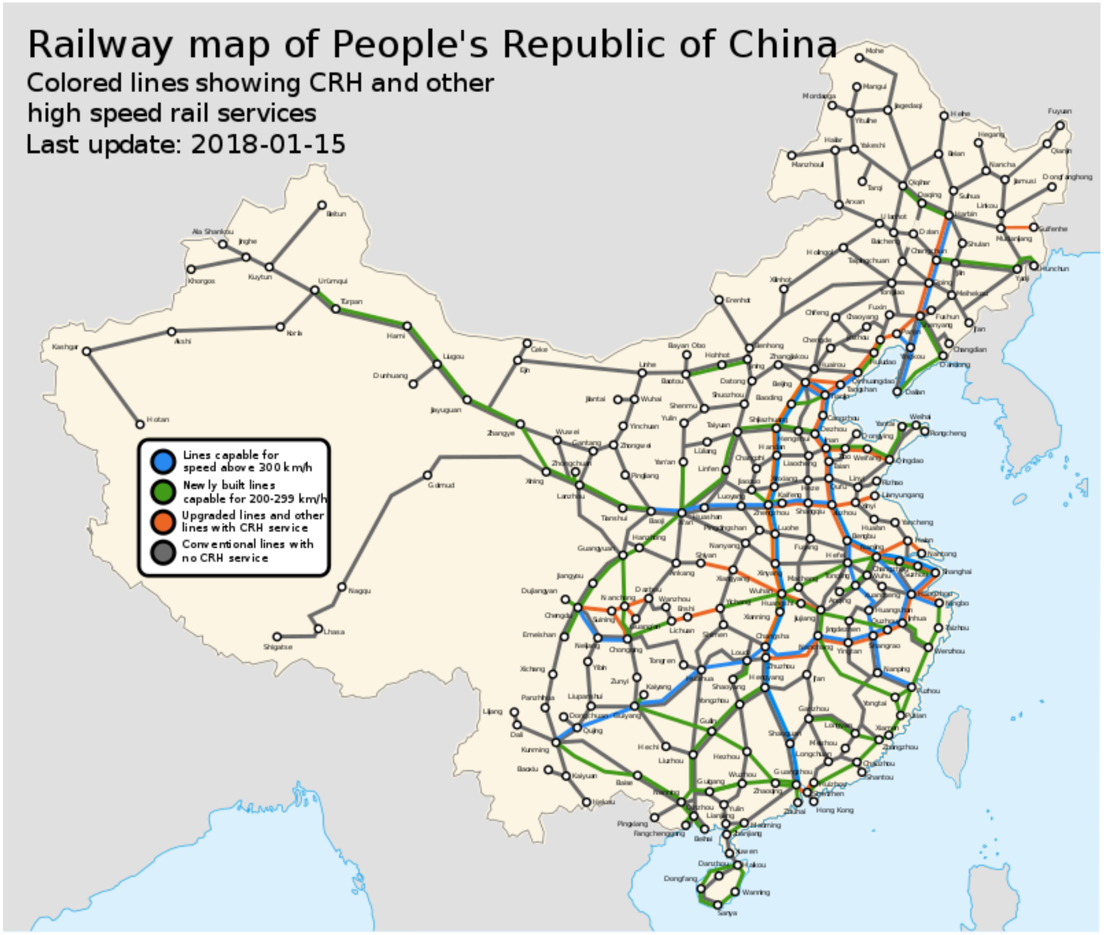

Figure 1.1: Railway Map in China

Definitions of HSR differ but, generally, railways with a maximum speed of 250 km/h or more are considered as HSR. China has built and operated a big number of High-Speed Railways and is making its high-speed rail network the longest in the world. As of end of the year 2017, China is operating HSR for 25,000 kilometers, which number accounts for 66.3% of the whole length all around the world. France is a good benchmark as its famous mature HSR network to understand the developing condition of China. If China was to become as many length of HSR line per inhabitant as France, its HSR network would reach 43,000 kilometers. According to the Mid-to-Long Term Railway Network Plan, which is adopted in 2004 first time by the state council of PRC, and newly updated in July 2016, China laid out a railway development

Table 1: Number of High-Speed Rail Stations over time (by the end of each year

	2010	2011	2012	2013	2014	2015	2016.04
Total Railway track length (km)	91,200	93,200	97,600	103,145	111,821	–	–
# of HSR stations	196	239	294	431	562	671	757
# of cities with HSR stations	68	79	99	114	157	175	175
# of cities with long-distance HSR stations	55	68	89	101	140	146	146
# of cities with intercity stations	21	21	21	28	41	63	66

Source: China Transportation and Communications Yearbook; China Railway Yearbook.

Figure 1.2: The No. of HSR Stations over Time

plan through 2025, including the connectivity of most cities above 500,000 people to a HSR network of 38,000 km, including 8 Verticals + 8 Horizontals Passageway Grid as the backbone, regional high-speed rail and high-speed intercity railways. Figure 1.2 and Figure 1.3 describe the HSR expansion from 2010 to 2015.

Mass start-up is the other economic wave recently. After 2010, the growth of Chinas economy slows down, labor price as a once main comparative advantage of Chinas export strategy has been increasing, and domestic consumer demands are still not enough high to support continuous development. Facing above economic pressures, China is trying to push a new supply-side structural reform, which is aimed at improve the endogenous power of its economy, to encourage people to start up new firms and innovate new business. Meanwhile, infrastructure investment such as HSR construction is still considered by Chinas government as a useful policy toolkit to facilitate new business entrepreneurship.

Although public infrastructure investment and private entrepreneurship both are important engines of economic growth, I do not know the quantitative and mechanical relationship between these two factors so far. Hence, I hope to detect the impact on firm entry (the number of newly registered firms) in county-level when the county is connected to the high-speed rail network, so as to measure the question I come up with.

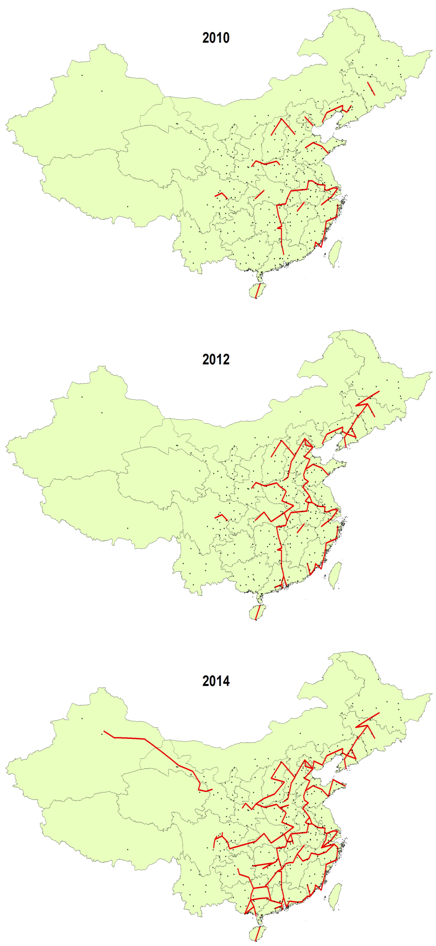

Figure 1.3: HSR Lines from 2010 to 2015

Chapter 2

Background & Literature

2.1 Firm Entry and Entrepreneurship

Entrepreneurship is the source of creative destruction, and is the way to make production combinations more efficient (Schumpeter 1934[14]). Hence, I can detect the condition of structural improvement of economy through entrepreneurship research. In our article, Firm entry can be regarded as an important economic index to label entrepreneurship.

Firm entry can be analyzed from regional dimension reasonably. Incessant firm entries to any regions are commonly arisen from the process of long-term economic growth, which is usually accompanied with urbanization and economic agglomeration. For example, the high-tech industry's booming in Silicon Valley benefited from horizontal integration of vibrant local small organizations, as opposed to inward looking tycoons in Route 128, Boston gradually lost their advantages compared to competitor SV (Saxenian 1996[13]).

Regional economic agglomeration can promote entrepreneurship. About the reasons of economic agglomeration, Marshall concluded with three micro-economic foundations: (1)Customer supplier linkages; (2)Labor market pooling; (3)Knowledge spillovers (Marshall 1920[10]). These economic forces make people settled together in urban areas. At the same time, some study argued that social interaction could be a parallel mechanism for promoting entrepreneurship in some economic cluster regions, including facilitative social network, social capital, and institutional structures (Porter 1998[11]). In concordance, some study showed that economic and social parts could correspond with each other from social network viewpoint, like customer supplier linkages are viewed as an inter-industry network defined by input-output relationship, labor market pooling is viewd as a network of occupational composition, and knowledge spillovers are viewed as using networks to distribute patents or knowledge within a given region (Cho 2018 [5]).

Here I know, firm entry phenomenon as a form of entrepreneurship, is related with regional economic and social interactions tightly.

2.2 Transportation Infrastructure and Firm Entry

Transportation infrastructure construction can reduce the cost of firm entry. As what I talked above, firm entry is geographically concentrated, which means that successful regions experience resurgences of agglomerative activity. Economic agglomeration and fragmentation are caused from the trade-off between increasing returns of scales and transaction costs, that include transportation and information costs. The costs are cheaper, economic activities for some regions and for some hierarchical organizations are more fragmentive. Transportation network can be considered as intermediate services in economic process of firms in different regions (Bell and Feitelson 1990[4]) The investments on transportation infrastructure can reduce the costs of service links, so that can promote vertical fragmentation which means that more firms can be started up, and promote horizontal agglomeration which means that the specific regions can attract more firms to enter (Jones and Kierzkowski 2001[7]).

Except the viewpoint of the specific city or company, the market network as a whole can also support our expectation. Work on network indicated that the aggregate characteristics of the whole network, except the specific nodes, ties or sub-groups, namely the density of production space is increased, like the HSR network in our case, the economic activities in the whole country level would also be improved (Hausmann and Klinger 2006[6]).

Lower cost to access intermediate inputs is another mechanism through improved network trade to promote total economic growth (Fujita, Krugman and Venables 1999[8]). Take other infrastructure like highway network in China as an example here, highway expansion affected regional GDP and population by reducing the costs of domestic trade, while accessing to export market is not important (Baum-Snow et al. 2016[3]).

Besides, transportation infrastructure can also promote the knowledge spillovers. Indian evidence showed that lower trade costs fostered innovation through competition (Topalova and Khandelwal 2011[15]). Roads evidence instead of railways showed that 10% increase of regional highways can causes 1.7% increase of regional patenting over a five-year period (Agrawal 2016[1]).

In conclusion, transportation infrastructure can lower the transaction cost including transportation, trade and information costs so that do help to more firms enter the whole and specific regional market.

2.3 Case of China

In the case of China, railways construction had been seen as the symbol of industrialization and national dignity. The ROC's railways system was taken over by the ministry of railways which was organized by Peoples Republic of China in 1949. In subsequent 60 years, this organization was responsible for the construction, operation and management of railways system under the central-planned economy. Gradually, the traffic capacity of old railways in the past time cannot support the rapid growth of Chinese economy after 21^{st} century. The ministry of railways activates a strategy of large-scale constructions and institution reforms in 2002, which is the start-up of the HSR network formation in China. As the reform goes deeply, the ministry was dismissed by the State Council, the function of administration was put under the ministry of transportation, and the function of operation and business was put under the new organized China Railway Corporation.

Since the HSR network expands from 2003 to now, the whole network of HSR improves 4.5% of the total welfare by reducing the travel costs and enhancing the connectivity among cities (Barwick et al. 2018[2]). In addition, face-to-face communications are very important for coordination and innovation which are the sources of new firms formation, and the HSR increases the city-level employment by 7% and passenger flows by 10% (Lin 2017[9]). Except to promote economic integration, HSR can also affect the distributional consequences. However, the affected counties would reduce 4-6% GDP relative to those non-affected ones, because investment likely to crowd into cities to pursuit higher increasing returns through HSR as a channel (Qin 201[12]).

Chapter 3

Data Description

3.1 Historical Trend

The data includes the number of newly registered firms, the number of operated HSR stations, and the number of railway lines in county-level for every month from 2011 to 2015. County and county-level division as the most basic local government is in the third level of administrative hierarchy under the prefecture-level city in most Provinces or Autonomous Regions, and the second level in Direct-Controlled Municipality. The data of newly registered firms could be recognized by 10 different industries, and the county could be ranked into 5 tiers by senior prefecture or municipality cities. To be mentioned, I just pick up five provinces, one autonomous region, and two municipality cities including Heilongjiang, Jilin, Liaoning, Inner Mongolia, Shanxi, Hebei, Beijing and Tianjin as research samples, and they are all located in the northeastern part of China.

I divided the counties into three classes according to the condition of HSR accessibility: (1) counties with no HSR line at the end of 2015, (2) counties with at least one HSR line operated at the beginning of 2011; (3) counties with their first line operated during 2011-2015. Based on this classification, I plot a figure to describe the historical trend of average firm entries for each class of counties in our time period.

As shown in the figure 3.1, the class with no HSR line operated during the period has the lowest level of average firm entries, the class with at least one HSR line operated before the period has the highest performance, and the the class that was connected to the HSR network for the first time in the period stay between that of the other two. In addition, all of them follow the fluctuations across time. Since January 2014, the number of newly firms in three classes of counties jumped up sharply and kept in a high level except in the end of the year 2014.

However, it is not clear if this classification can show the causation between HSR connectivity and firm entries because a county with greater economic behavior and the regions with more firm entries are more likely to be connected to HSR network earlier, or it supports that connecting to HSR has a positive impact on firm entries, and the earlier the operation happens or the city tier is higher, the stronger the positive impact could be.

3.2 Patterns Before and After HSR Opening

As our focus is the impact of first HSR operated station on the average firm entries, not the other way around, I choose to set the operation time as a cut point for each county to see if the regression lines look different before and after the cut point, and if there is a discontinuity gap.

I plot the average firm entries across time for the two years (24 months) before operation and the two years (24 months) after operation. I only use the data for the third class, the counties with their first line operated during 2011-2015, because I only know the exact operation time information for this class. As shown in the figure

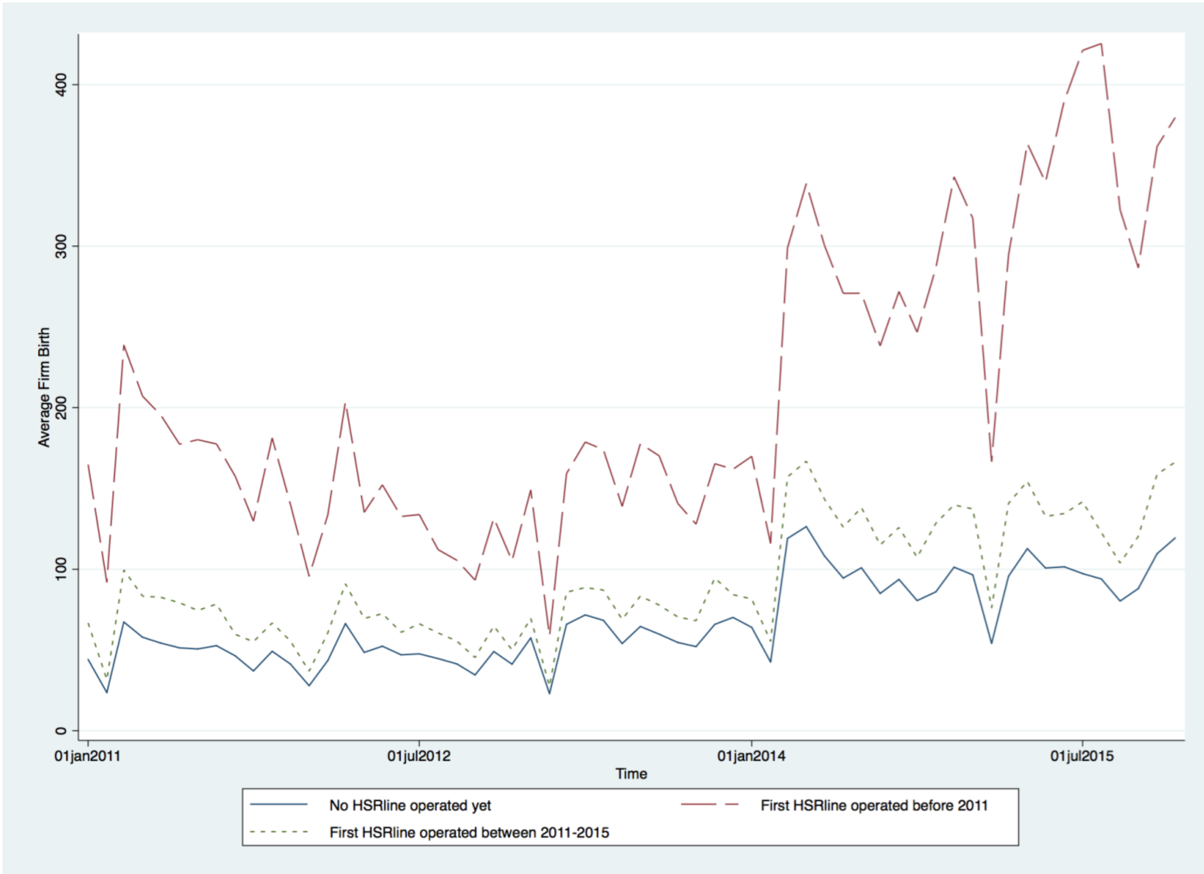

Figure 3.1: Historical Average Firm Entry

3.2, theres a jump at the cut point (the month the HSR line started operating). It shows the operation of the HSR has a positive relationship on average firm entries.

However, I am still not sure if the jump is caused by the increasing trend of firm entry across this time period or the characteristics of the counties chose to build HSR line. So, the next step is to exclude these two effects, namely time and region fixed effects.

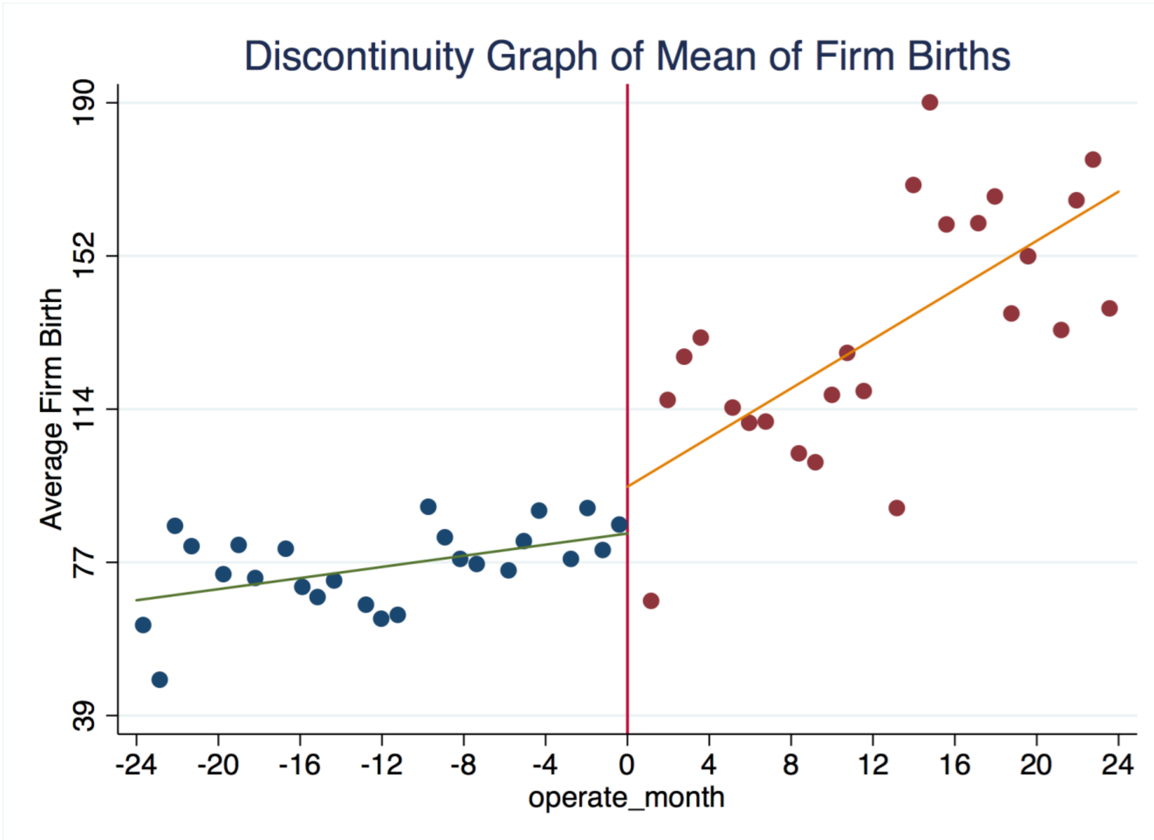

Figure 3.2: Discontinuity Pattern

3.3 Residuals Discontinuity with Fixed Effects

In order to exclude the fixed effect of time and county, I run the following regression to get the residuals:

$$Log(FirmEntry)_{it} = \beta_0 + \gamma_1 month_t + \gamma_2 county_i + \epsilon_{it} \tag{3.1}$$

Where the dependent variable is the log of number of firm entries for each county i in time t. Variable *month* denotes the fixed effect of time, and variable *county* denotes the fixed effect of region. By using the residuals to draw the graph to show the difference before and after the first station operated, I exclude the effects bring by the overall time trend and the counties characteristics. As shown in the figure 3.3,

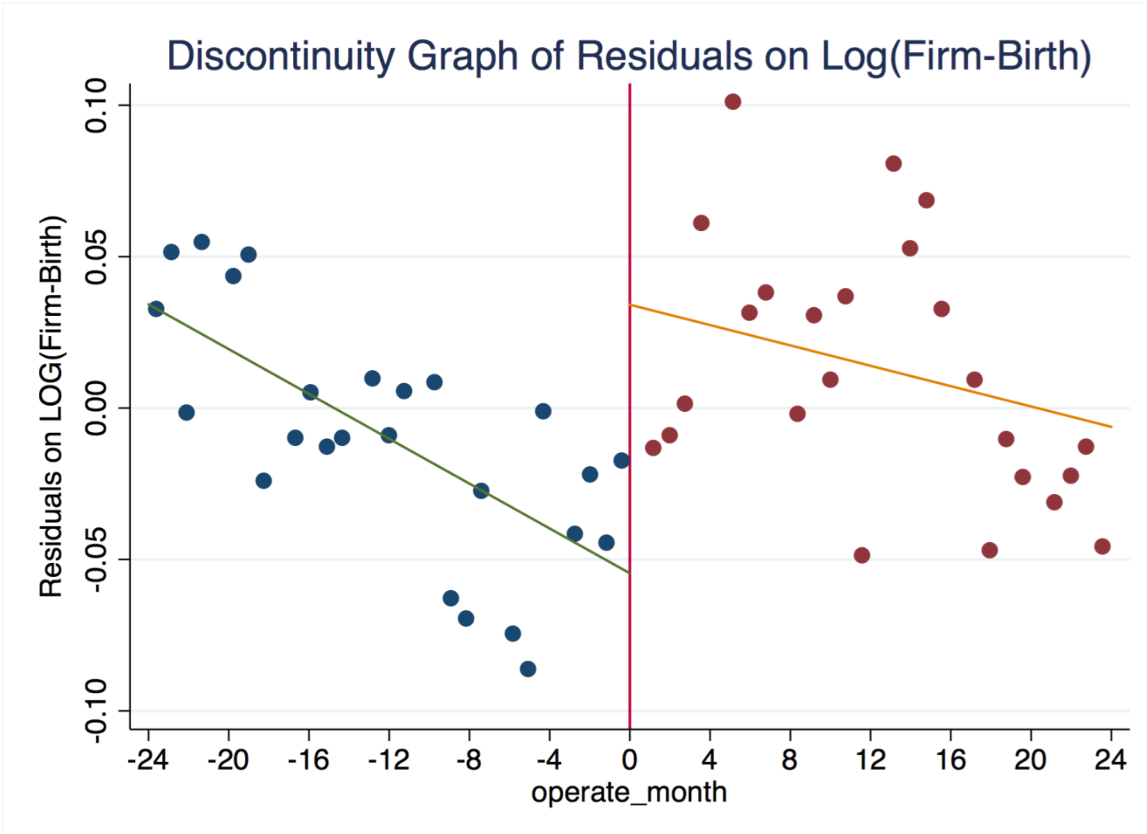

Figure 3.3: Discontinuity with Fixed Effects

the jump at the cut point becomes even more apparent and supports our hypothesis that being connected to the HSR network can increase the firm entries in general.

However, this estimation still can not exclude the influences of county specific time trend. I would add this control variable to exclude its effects.

3.4 Residuals Discontinuity with Fixed Effects and County Specific Time Trend

At this step, I add the fixed effect of time, county and county specific time trend controller to the test the equation. I run the following regression to get the residuals:

$$Log(FirmEntry)_{it} = \beta_0 + \gamma_1 month_t + \gamma_2 county_i + \gamma_3 county_i \times month_t + \epsilon_{it} \quad (3.2)$$

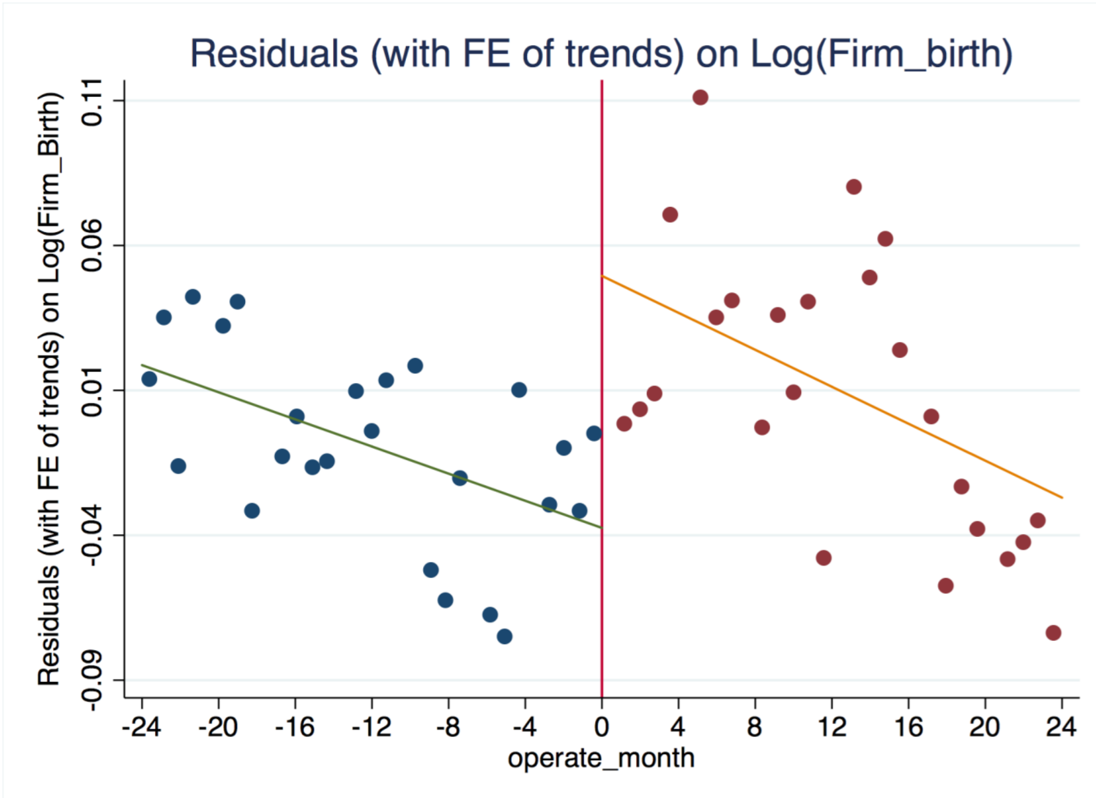

Figure 3.4: Discontinuity with Fixed Effects and County Specific Time Trend

Where control variable $county \times month$ denotes the the county specific time trend. By using the residuals to draw the graph again, I exclude above three effects. As shown in the figure 3.4, the jump at the cut point does not shift too much compared with the last one. Here I know after excluding other possible effects, there is still an apparent gap at the cut point time, which means that HSR station's operation may do positive influence on firm entry in our sample regions. Next I would build a model to get a quantitative result.

Table 3.1: Summary Statistics in Sample Cities in 2015

City Name	# of Counties	# of Lines Built	# of Lines operated	Annual Firm Entry	Average Firm Capital	Growth Rate Since 2011
Beijing	14	7	4	190883	1529710	91%
Tianjin	11	11	8	39250	237426.8	78%
Shijiazhuang	24	10	6	44482	91221.28	166%
Tangshan	11	2	2	9146	63368.02	137%
Qinhuangdao	8	4	4	7562	34065.28	117%
Handan	20	1	1	20312	32124.85	178%
Xingtai	20	1	1	14782	24208.39	171%
Baoding	22	6	3	16724	27477.94	198%
Cangzhou	17	1	1	14189	26122.94	164%
Langfang	11	2	1	13733	55681.07	220%
Taiyuan	10	4	4	13799	66069.36	92%
Datong	12	0	0	4053	12832.16	46%
Yangquan	5	1	1	2012	11989.81	72%
Changzhi	13	0	0	4717	12287.02	12%
Jinzhong	12	6	6	4795	19291.07	77%
Yuncheng	13	3	3	7540	14363.44	107%
Linfen	18	5	5	6403	11583.81	81%
Huhehaote	10	0	0	12015	62077.27	185%
Baotou	10	0	0	6564	36681.44	45%
Wuhai	3	0	0	1343	12106.7	109%
Chifeng	13	0	0	8674	20879.23	202%
Tongliao	8	0	0	5222	25387.84	213%
Eerduosi	8	0	0	6414	34899.73	4%
Shenyang	13	12	6	27587	97697.95	53%
Dalian	10	9	4	24610	209146.8	83%
Anshan	7	4	4	4341	25649.78	31%
Fushun	8	0	0	3654	19548.54	53%
Dandong	7	5	1	4323	19029.66	-15%
Jinzhou	6	1	0	4287	31681.77	92%
Yingkou	7	3	3	4795	41632.88	18%
Liaoyang	7	1	1	1869	23555.03	34%
Panjin	5	3	3	2873	34688.93	50%
Huludao	7	3	3	4148	27160.34	64%
Changchun	11	5	5	23549	130471.7	90%
Jilin	9	3	3	5746	12761.8	77%
Haerbin	17	4	3	22058	60217.26	100%
Qiqihaer	16	1	1	5123	11393.14	59%
Daqing	9	3	1	4225	25065.57	111%
Mudanjiang	10	0	0	4507	14929.03	99%

Chapter 4

Empirical Model

In the model, I consider the impacts of HSR station operated on the firm entry for each county. I divide all counties into two groups of cities based on city tiers first. The ones are big cities which are belong to tier 1, 2 and 3, the other ones are small cities which are belong to tier 4 and 5. Then I run three regressions with different dependent variables. The first one is on the whole counties, the second one is on the big cities, and the third one is on the small cities.

$$Log(FirmEntry)_{it} = \beta_0+\beta_1 HSR_i \times Post_t+\gamma_1 County_i+\gamma_2 Month_t+\gamma_3 Month_t \times County_i+\epsilon_{it} \tag{4.1}$$

HSR $\times$ *Post* is a difference in difference term, where the dummy variable *HSR* denotes whether county i has any station operated in 2011-2015 or not (=1/0 if operates/not operates); and *After* denotes whether it is after the station operated for each time t or not. So the DiD term denotes for time t and county i whether there is a operated station or not. *County* denotes the region fixed effect, which control the different characteristics of different counties, like administrative hierarchy, culture, economic and social development. *Month* denotes the time fixed effect, which control the impacts of economic developing trend on firm entries, because the firm entry would also be influenced by economic growth or decline across the passed time. *Month* $\times$ *County* denotes a control variable that control the different time trends of different counties, like county A and county B they may have different GDP growth rates. In the dependent variables' side, I do a logarithm of number of the firm entry for whole sample, big cities and small cities, which means to calculate the percentage change of number of the firm entry bring by the operated HSR station.

Table 4.1: Impacts of HSR on firm entry from OLS

	(1)	(2)	(3)
VARIABLES	Whole Sample	Citytier 1&2&3	Citytier 4&5
HSR*Post	0.0572***	0.0323	0.0720***
	(0.0182)	(0.0288)	(0.0233)
Constant	3.186***	3.858***	2.877***
	(0.0190)	(0.0327)	(0.0232)
Time FE	YES	YES	YES
County FE	YES	YES	YES
Time Trend	YES	YES	YES
Observations	26,438	8,338	18,100
R-squared	0.618	0.599	0.626
Number of id	451	141	310

Notes: Outcome variables are ln number of firms in each county. Our regression controls for year-month fixed effects, county level fixed effect, City$\times$month fixed effect, and city-specific linear time trends. Standard errors are clustered at the county level. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Based on our result table 4.1, the operated HSR station would do very significant impacts on counties' firm entries at 99% level for all the sample. In the first column, I can see the parameter is 0.0572, which means that once a county is connected to HSR network, there would be 5.72% increase of newly registered firms in this county. Unfortunately, the impacts on big cities are not significant, which indicates that HSR station would not do important helps on local firm entry. However, operated station can still bring significant positive influence on small cities in the third column. I find once a county which belongs to small city is connected to HSR network, there would be 7.20% increase of newly registered firms at 99% level significance.

Here I can make a conclusion, HSR connectivity can do significant positive impact on regional firm entry, especially on small cities, but on big cities the impact is insignificant. Hence, HSR network construction can do help to entrepreneurship of the whole regions of this part of China, and also do help to the rebalance of regional development imparity.

As I know, agglomeration is one of the sources of modern economic growth and entrepreneurship. HSR network construction improve the market network density, save the transportation, communication and transaction costs, facilitate the spillovers of knowledge across the regions, so as to improve the firm entry level of the whole regions.

HSR construction can lower the transportation and information cost among economic subjects. Economic activities would be more fragmentative and distributed more evenly across the lands. For small cities, HSR construction can help them to access a bigger market, so that more firms would enter these regions to enjoy lower land cost and to get the benefit of transportation welfare comes from HSR network. For firms in big cities, HSR construction facilitate them to fragment new production branches to locate in other small cities. For big cities, more mature HSR network weaken the comparative advantages of transaction compared to other cities, and they have more expensive land cost at the same time, so they would not get extra positive benefit from HSR system.

Chapter 5

Robustness Checks

5.1 Excluding counties that were treated before 2011

I exclude the counties that have had operated stations before 2011, namely I just keep counties with ***new*** HSR stations during the time period. There are 429 counties in the whole sample, 128 counties that belong to big cities(tier 1&2&3), and 301 counties that belong to small cities(tier 4&5) right now. Then I rerun the model.

I do this robustness check because there might be different impacts between the first operated station and the additional operated station. If the county has been treated before 2011, then the treatment would comes from a additional HSR station but not a new one. To control this influence, I exclude the counties that have had operated stations before 2011, so that the rest counties are all treated first time during 2011-2015.

Based on the result of the table 5.1, I find that the result is similar with our original model in table x.x. In column (1), the impact on the whole sample is significant at 99% level, and new operated HSR station would bring the county with 5.45% growth on firm entry. It is a bit lower than 5.72% of last model but not too much. The

Table 5.1: Impacts of HSR on firm entry from OLS - excluding counties treated before 2011

	(1)	(2)	(3)
VARIABLES	Whole Sample	Citytier 1&2&3	Citytier 4&5
HSR*Post	0.0545***	0.0257	0.0715***
	(0.0185)	(0.0297)	(0.0236)
Constant	3.140***	3.776***	2.867***
	(0.0196)	(0.0350)	(0.0237)
Time FE	YES	YES	YES
County FE	YES	YES	YES
City Time Trend	YES	YES	YES
Observations	25,118	7,558	17,560
R-squared	0.618	0.594	0.627
Number of id	429	128	301

Notes: Outcome variables are ln number of firms in each county. Our regression controls for year-month fixed effects, county level fixed effect, and city-specific linear time trends. Standard errors are clustered at the county level. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

impacts on big cities are still not significant, which indicates that even new HSR station would not do important helps on local firm entry. New operated station can still bring significant positive influence on small cities in the column (3). I find once a county which belongs to small city is connected to HSR network first time, there would be 7.15% increase of newly registered firms at 99% level significance. This result is very close to the original 7.20%. Small cities might easily be influenced by outside treatment, so new HSR operated station might give more helps to small cities, but the differences with original model are really small. The finding from this robustness check supports that our model's result is robust.

5.2 Excluding counties without HSR stations

I exclude the counties without HSR station as of 2015, namely just keep counties with HSR stations during the time period. There are 80 counties in this sub-sample, 34 counties that belong to big cities(tier 1&2&3), and 46 counties that belong to small cities(tier 4&5).

I do this robustness check because the HSR station location is not a random behavior because the station location is decided by government. Government may likely to choose the location where more firms would like to enter. To avoid this possible endogeneous problem, I exclude the counties without stations to mitigate the endogeneity concern.

Table 5.2: Impacts of HSR on firm entry from OLS - counties with HSR stations

	(1)	(2)	(3)
VARIABLES	Whole Sample	Citytier 1&2&3	Citytier 4&5
$HSR \times Post$	0.0596**	0.0633**	0.0488
	(0.0282)	(0.0309)	(0.0583)
Constant	3.784***	4.416***	3.309***
	(0.0340)	(0.0492)	(0.0475)
Time FE	YES	YES	YES
County FE	YES	YES	YES
City Time Trend	YES	YES	YES
Observations	4,749	2,040	2,709
R-squared	0.795	0.793	0.798
Number of Counties	80	34	46

Notes: Outcome variables are ln number of firms in each county. Our regression controls for year-month fixed effects, county level fixed effect, and city-specific linear time trends. I only include the counties with HSR stations. Standard errors are clustered at the county level. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Based on the result of the table 5.2, I find that the impact of station operated on all the counties with stations is significant at 99%. Once the station is operated, the firm entry would increase by 5.96%, which is a bit higher but still similar to the original model result 5.72%. However, the impacts on big and small cities are totally different, and the results of significance are changed with each other. In column (2), I find the result of big cities become significant on 95% level, and operated HSR station would bring the county 6.33% growth on firm entry. In column (3), the impact on small cities is not significant anymore. This could be driven by the much smaller sample and hence larger standard errors in this analysis.

We get a reverse result with original model about impacts on city tiers. In this check, we exclude the influences of human decision about station location. Then

HSR connectivity has more positive impacts on big cities. However, in column (2) and (3), the standard errors are both big. The one for big cities is 0.0309, and the other one for small cities is 0.0583, so the two regression distributions of coefficient $HSR \times Post$ are very diverse, and share a big part together. Hence, I cannot get a robust conclusion of the diffrent impacts between big cities and small cities.

5.3 Impacts on Counties in HSR-connected cities

The county without operated station could also be influenced by HSR accessibility if its senior city is connected to the HSR network. In this check, I exclude all the cities without operated station, which means all the samples are the subordinate counties of the cities connected with HSR system. I have 349 counties of the whole sample, 121 counties that belong to big cities(tier1&2&3), and 228 counties that belong to small cities(tier 4&5).

Table 5.3: Impacts of HSR on firm entry from OLS - counties in HSR-connected cities

VARIABLES	(1) Whole Sample	(2) Citytier 1&2&3	(3) Citytier 4&5
$HSR \times Post$	0.0572***	0.0323	0.0720***
	(0.0173)	(0.0259)	(0.0229)
Constant	3.306***	3.994***	2.934***
	(0.0205)	(0.0316)	(0.0266)
Time FE	YES	YES	YES
County FE	YES	YES	YES
City Time Trend	YES	YES	YES
Observations	20,446	7,183	13,263
R-squared	0.632	0.643	0.627
Number of Counties	349	121	228

Notes: Outcome variables are ln number of firms in each county. Our regression controls for year-month fixed effects, county level fixed effect, and city-specific linear time trends. I only include the counties located in HSR-connected cities. Standard errors are clustered at the county level. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Based on the result of table 5.3, I find that the impact on whole sample is the same with original model, and it is also 5.72% at 99% significant. In column (2), the impact on big city are still insignificant. In column (3), for small cities, the impact is also the same with original model, and it is 7.2% at 99% significant. The reason I get almost the same result is what samples I drop are no any station, and what I keep are almost the same with our original model.

This check can support that our model's result is robust.

5.4 Heterogeneous Impacts by Industry

Finally, I do a regression on counties with different industries by using original model. I divide firms sample into ten different industries, transportation, resident service, real estate, information & software, construction, scientific research & integrated tech, agriculture & forestry, leasing & business service, manufacturing, and whole sale & retail trade. I put these ten types of industries samples as dependent variables separately instead of the whole sample of firm entry. I run the regression and draw the following figure 5.1 then. In the figure, red bars are the results of which industries can get significant influences from HSR station operation, and blue bars are the results of which industries can get insignificant influences from station operation. The vertical axis reflects the coefficient number of treated variable $HSR\times$ Post.

Based on figure 5.1, I find that information & software, scientific research & integrated tech, leasing & business service, manufacturing, wholesale & retail trade, which industries are impacted by operated station significantly, and transportation, resident service, real estate, construction and agriculture & forestry, which industries are impacted by operated station insignificantly. The interesting finding is that the insignificant industries are more land intensive, and significant industries are more labor or knowledge intensive. Chinas HSR is a system for human transportation, more operated HSR on regions can improve the ability of local industry, especially the industries that require human mobility, to access outside market. Leasing and business

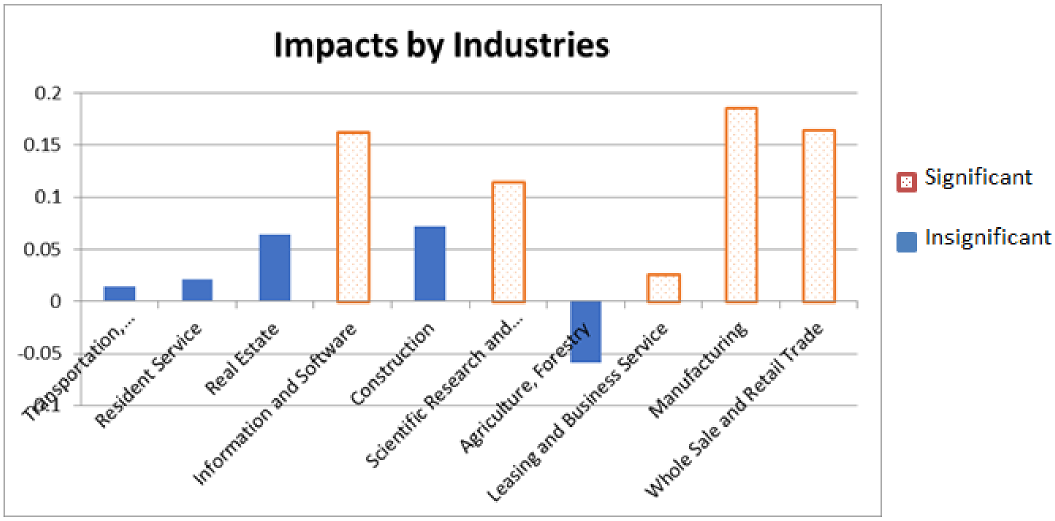

Figure 5.1: Impacts by Industry

service is relatively influenced slightly in five significant industries, because which is a industry that relatively dependent on the coordination services inside the cities but not mobile between regions too much. Manufacturing, whole sale and retail trade have the highest influences coming from HSR station operation, which are highly labor intensive industries. Information, software, scientific research and integrated tech have the second high influences coming from HSR station operation, which are highly knowledge intensive industries. The former one benefit from the lower trade and transportation cost, and the latter one benefit from the more knowledge spillover. This finding is consistent with our intuition and previous researches.

Chapter 6

Conclusion

In this thesis, I carry out an empirical analysis to examine the quantitative impact of HSR expansion on firm entry using county-level firm registration data in China from 2011 to 2015. The difference-in-differences analysis leads to three findings.

First, the availability of HSR stations in a county is associated with a 5.72% increase in the number of new firm registration in that country on average. This finding is robust to different specifications including the choice of control groups. This suggests that transportation infrastructure investment could spur firm activities and promote entrepreneurship. The positive impact on firm entry could be driven by the more frequent exchanges of ideas and technology know-how across regions.

Second, the baseline specification and the specification excluding counties with HSR before our data period both suggest that the positive impact of HSR expansion on firm entry is larger in smaller cities (tiers 4 and 5). This implies that transportation infrastructure investment could serve to reduce regional economic inequality in China.

The data in my study covers the northeastern part of China and significant differences exist within this region. Beijing serves as the cultural, educational, political, industrial and commercial center of China. It attracts top talents from all over China especially the northern part of China. Tianjin, next to Beijing is a large port city and is also an important industrial center. Other provinces and autonomous region in this region are relatively less developed compared to Beijing and Tianjin. Many are facing significant challenges in the transition from manufacturing-based economy to service-based economy. My analysis provides suggestive evidence that HSR expansion could help reduce economic imbalance in this region by promoting new entry of firms and economic growth in the less developed areas.

Third, the analysis by industry shows that the positive impact on firm entry is larger among knowledge or information-intensive industries such as research, information and software. The impact on industries that rely heavily on fixed production factor (e.g., land) including agricultural and real estate sectors is smaller or non-existent. The HSR system transports passengers rather than freight and the more frequent passenger travels could help the exchange of ideas and promote entrepreneurial activities especially in the industries that rely more heavily on human capital.

There are several directions for future research. First, it would be interesting to expand my analysis on eight provinces in the northern part of China to the national level as other regions such as the east and south have different economic and social characteristics compared to the study area in this research. Second, it would be important to further understand the channels of impacts and the heterogeneity in impacts across cities and by industry. Firm-level data and information on passenger flows should be helpful for this line of inquiry. Third, the broader impacts o HSR expansion on the migration of firms and people and ultimately on economic performance across regions are worth future research.

Bibliography

[1] Ajay Agrawal, Alberto Galasso, and Alexander Oettl. Roads and innovation. *Review of Economics and Statistics*, 99(3):417–434, 2017.

[2] Panle Jia Barwick, Dave Donaldson, Shanjun Li, and Yatang Lin. The welfare effects of passenger transportation infrastructure: Evidence from chinas hsr network. 2018.

[3] Nathaniel Baum-Snow, Vernon Henderson, Matthew A Turner, Qinghua Zhang, and Loren Brandt. Highways, market access, and urban growth in china. *SERC Discussion Papers*, 200, 2016.

[4] Michael Bell and Eran Feitelson. Bottlenecks and flexibility: Key concepts for identifying economic development impacts of transportation services. *Transportation Research Record*, 1274:53–59, 1990.

[5] Jae Beum Cho. Agglomeration, regional social capital, and entrepreneurship. 2018.

[6] Ricardo Hausmann and Bailey Klinger. Structural transformation and patterns of comparative advantage in the product space. 2006.

[7] Ronald W Jones and Henryk Kierzkowski. Horizontal aspects of vertical fragmentation. In *Global production and trade in East Asia*, pages 33–51. Springer, 2001.

[8] Paul Krugman, Masahisa Fujita, and Anthony Venables. The spatial economy. *Cambridge: MIT Press*, 1:999, 1999.

[9] Yatang Lin. Travel costs and urban specialization patterns: Evidence from chinas high speed railway system. *Journal of Urban Economics*, 98:98–123, 2017.

[10] Alfred Marshall and Mary Paley Marshall. *The economics of industry*. Macmillan and Company, 1920.

[11] Michael E Porter. *Clusters and the new economics of competition*, volume 76. Harvard Business Review Boston, 1998.

[12] Yu Qin. no county left behind?the distributional impact of high-speed rail upgrades in china. *Journal of Economic Geography*, 17(3):489–520, 2017.

[13] AnnaLee Saxenian. *Regional advantage*. Harvard University Press, 1996.

[14] Joseph A Schumpeter. Change and the entrepreneur. *Essays of JA Schumpeter*, 1934.

[15] Petia Topalova and Amit Khandelwal. Trade liberalization and firm productivity: The case of india. *Review of economics and statistics*, 93(3):995–1009, 2011.